

Metaanálisis, megaensayos y práctica clínica en cardiología

Carlos Brotons Cuixart y Gaietà Permanyer Miralda

Unidad de Epidemiología Clínica. Servicio de Cardiología.
Hospital Universitario Vall d'Hebron. Barcelona.

ensayos clínicos / evaluación de procesos y resultados / intervalos de confianza / metaanálisis / mortalidad / prácticas clínicas / tamaño muestral

La agrupación de los datos de diferentes ensayos clínicos es el factor que crea más escepticismo y preocupación en los médicos clínicos. Para poder agrupar los resultados de diferentes ensayos clínicos es preciso que se cumplan tres requisitos: similitud en los tratamientos, similitud de pacientes y similitud en las variables de resultado.

El sesgo de publicación es una de las limitaciones de los metaanálisis y se produce por la no publicación de los estudios que dan resultados negativos, pudiendo sobrevalorar el efecto observado en el metaanálisis.

Podría parecer que un efecto demostrado para el conjunto de la población de estudio con un tamaño de muestra grande se podría analizar fácilmente en diferentes subgrupos de pacientes y acercarnos así a la predicción del efecto en el paciente individual. Esta observación aparentemente obvia es una falacia y el análisis de subgrupos a menudo es problemático. Posiblemente la condición más importante del análisis de subgrupos es la definición de los subgrupos en la fase de diseño del estudio, cosa que puede permitir una previsión adecuada de los requisitos necesarios para la validez. El problema de generalización de los resultados tradicionalmente se ha relacionado con el hecho de extrapolar los mismos a grupos de pacientes no incluidos en el ensayo clínico. En estos casos es importante considerar otros estudios epidemiológicos no experimentales, su validez interna y la consistencia entre los mismos.

Las características fundamentales que diferencian los ensayos clínicos convencionales de los megaensayos son las siguientes: los megaensayos reclutan una población muy grande y heterogénea, con unos criterios de inclusión y exclusión poco estrictos; la única variable de resultado que se puede determinar mejor en estas condiciones tan heterogéneas es la mortalidad; debido al tamaño tan grande de la muestra sus resultados consiguen un alto grado de precisión (intervalos de confianza muy estrechos).

Palabras clave: *Metaanálisis. Megaensayos. Práctica clínica.*

Correspondencia: Dr. C. Brotons Cuixart.
Unidad de Epidemiología Clínica.
Servicio de Cardiología. Hospitals Vall d'Hebron.
P.º Vall d'Hebron, 119-129. 08035 Barcelona.
Correo electrónico: brotons@hg.vhebron.es

META-ANALYSIS, MEGATRIALS AND CLINICAL PRACTICE IN CARDIOLOGY

One of the factors that raises more scepticism and concern in physicians is the combination of data from different clinical trials. In order to combine the results of different trials three requisites must be followed through: similar treatment, the patients have the same disease, and the main outcome variables must be the same.

Publication bias is one important limitation of meta-analysis, and it occurs when studies with negative results are not published, which causes the effect of the treatment to be overestimated. It might seem reasonable that the demonstrated effect in the entire study population with a large sample size could be easily analyzed in different subgroups of patients, getting closer to the prediction of effect in the individual patient. This apparently obvious observation is a fallacy and subgroup analysis is often problematic. Possibly the most important condition in the analysis of subgroups is the definition of the subgroups in the design stage of the study, which can permit an adequate presight of the necessary requisites for the validity of the study. The problem of generalization of the results has been traditionally related with of extrapolation of the results to certain groups of patients not included in the clinical trial. In these cases it is important to consider other nonexperimental epidemiological studies, its internal validity and the consistency of their results.

The most important characteristics that differentiate conventional clinical trials from megatrials are the following: megatrials recruit a very large and heterogenous population, with few inclusion and exclusion criteria; the only outcome variable that can be better assessed in these heterogenous conditions is eventually mortality; due to their very large sample size, results are obtained with a high level of precision (very narrow confidence intervals).

Key words: *Meta-analysis. Megatrials. Clinical practice.*

(*Rev Esp Cardiol* 1999; 52: 840-850)

INTRODUCCIÓN

La introducción de los ensayos clínicos con asignación aleatoria representa uno de los avances fundamentales en la terapéutica y la investigación clínica de la segunda mitad del siglo xx. Los metaanálisis y los megaensayos se han desarrollado con posterioridad y han permitido demostrar, en algunas ocasiones, la eficacia de una intervención terapéutica determinada que no hubiera estado fácilmente demostrada con uno o varios ensayos clínicos de tamaño pequeño. Otras veces, incluso, han permitido refutar algunas hipótesis previas basadas en ensayos clínicos también de tamaño más pequeño. La característica común de los metaanálisis y los megaensayos es que aleatorizan poblaciones de tamaño muy grande (en el primero a partir de la agrupación de los datos de diferentes estudios, y en el segundo, a partir de un único estudio).

Este artículo no pretende revisar los aspectos formales y metodológicos de los metaanálisis y de los megaensayos, porque ya existe abundante bibliografía al respecto. Por ejemplo, recientemente, en esta misma revista, se publicó un excelente artículo sobre metaanálisis y revisiones sistemáticas en cardiología¹. Un tema como éste, que ha experimentado un crecimiento exponencial en las referencias bibliográficas (de menos de 50 citas en 1987 a casi 800 en 1996 con el encabezamiento de metaanálisis), y que ha aportado resultados relevantes para la práctica clínica en temas como el tratamiento del infarto agudo de miocardio o la prevención cardiovascular, pensamos que merece una reflexión destacando algunos aspectos metodológicos, las incertidumbres de la aplicación directa de sus resultados en la práctica clínica, y líneas de investigación futuras que a nuestro entender podrían aportar conocimientos valiosos respecto a algunas lagunas todavía hoy día existentes. Algunos de estos aspectos se discutirán conjuntamente para los metaanálisis y megaensayos, y otros se harán por separado, empezando por el metaanálisis.

¿QUÉ ES UN METAANÁLISIS?

Es un análisis estadístico que combina o integra los resultados de diferentes ensayos clínicos independientes que son considerados «combinables». El término de revisión sistemática es más apropiado, es decir, una descripción extensa y completa de la evidencia existente en un tema particular teniendo en cuenta todas sus virtudes y sus limitaciones. De hecho, las denominadas revistas secundarias como *Evidence-Based Medicine* o *Evidence Based-Medicine in Cardiovascular Diseases*, revistas que comentan críticamente artículos originales de la bibliografía médica, hablan de Revisión («Review» u «Overview») en vez de metaanálisis. La revisión sistemática consta básicamente de dos componentes: uno cualitativo, donde se hace una des-

cripción epidemiológica de los estudios, actuando los estudios individuales como los sujetos de la investigación, y un componente cuantitativo, que es el metaanálisis propiamente dicho, y que se refiere a la agrupación estadística de los resultados, obteniéndose un estimador de efecto único, que combina los efectos observados en los diferentes estudios individuales. Sin duda este último componente, que se refiere puramente a la técnica estadística, es el que tiene más impacto en el colectivo médico y el que refleja de una manera más directa el efecto de una determinada intervención. Sería interesante que los autores de los metaanálisis incluyeran siempre al menos una tabla en la que se describieran las características de los diferentes estudios, sobre todo las relacionadas con su validez (fuente de los datos, las variables de confusión que se midieron, las variables principales y secundarias de resultado, las pérdidas en el seguimiento y otras). Pudiera muy bien ocurrir que en una revisión sistemática no se realizara el componente cuantitativo, es decir, el metaanálisis, pero siguiera aportando datos de interés tanto metodológicos como clínicos. De hecho, una de las mayores críticas al metaanálisis es que se le da demasiado valor a las conclusiones cuantitativas, para a veces, evitar entrar en los resultados cualitativos, que podrían reflejar una agrupación de estudios inadecuada o una mala calidad de los estudios.

A lo largo de este artículo seguiremos hablando de metaanálisis, sobre todo porque en los múltiples ejemplos que se comentarán publicados en la bibliografía, y más en los artículos menos recientes, en casi todos se utiliza esta terminología.

¿CUÁNDO SE PUEDEN AGRUPAR LOS RESULTADOS DE LOS ENSAYOS CLÍNICOS?

La agrupación de los datos de diferentes ensayos clínicos es el factor que crea más escepticismo y preocupación, y no sin razón, en los médicos clínicos. Si el proceso de agrupación de los datos anula inadvertidamente alguna evidencia clínicamente importante de alguno de los ensayos, entonces el metaanálisis puede producir más daño que beneficio en la práctica clínica. Para poder agrupar los resultados de diferentes ensayos clínicos es preciso que se cumplan tres requisitos (asumiendo que todos se plantean la misma pregunta de investigación):

1. Similitud en los tratamientos: uno de los primeros metaanálisis en patología cardiovascular evaluó el efecto de los betabloqueantes en la mortalidad total en pacientes que habían padecido un infarto de miocardio². En este estudio se consideró que todos los betabloqueantes incluidos en los diferentes estudios eran equivalentes a efectos clínicos. También se realizaron dos análisis agrupados, dependiendo de si el trata-

miento se había empezado justo después del infarto o al cabo de unos días.

2. Similitud de pacientes o participantes en el ensayo: el metaanálisis que evaluó el efecto de los antiagregantes plaquetarios en la prevención de nuevos eventos vasculares incluyó ensayos con pacientes con enfermedad cerebrovascular, infarto de miocardio o angina inestable³. En este estudio, y de una manera razonable, el primer paso que se realizó fue agrupar los resultados por enfermedades y obtener los estimadores de efecto para cada enfermedad cardiovascular. Después se obtuvo un estimador global, que aunque agrupaba todas las enfermedades cardiovasculares tenía menos sentido clínico o interpretación más cuestionable que los anteriores.

3. Similitud en las variables de resultado: si se trata de mortalidad cardiovascular o eventos mayores cardiovasculares no suele haber serios problemas de definición o validación del resultado. Sí es importante diferenciar los tiempos en que se producen los eventos, como en el caso de la trombólisis en el tratamiento del infarto de miocardio, en la que es importante distinguir si es mortalidad a corto plazo o a largo plazo.

Si existe gran heterogeneidad clínica entre los estudios, y concretamente en estos tres aspectos comentados, difícilmente se podrán agrupar para realizar un metaanálisis. También podemos observar heterogeneidad cuando los estimadores de efecto de los diferentes estudios individuales van en direcciones opuestas, o que un estudio no encuentre un efecto y otros encuentren un efecto de magnitud muy importante. La presencia o no de heterogeneidad se puede analizar con tests estadísticos y/o con métodos gráficos. Se trata de un ejercicio matemático, y como tal, le corresponde hacerlo a un estadístico. La falta de heterogeneidad estadística no se puede considerar, sin embargo, como evidencia de homogeneidad. Y esto es así porque no sólo una prueba no significativa no se puede considerar como evidencia directa a favor de la hipótesis nula, sino también porque las pruebas estadísticas de heterogeneidad tienen en general una potencia baja y pueden dejar de detectar diferencias estadísticamente significativas incluso delante de niveles moderados de heterogeneidad genuina. Para poder explicar la heterogeneidad se necesita imaginación, sentido común, experiencia clínica y experiencia en investigación. Nunca deberían combinarse los resultados de los diferentes ensayos clínicos sin antes haber analizado y descartado la presencia de heterogeneidad clínica y estadística.

ALGUNAS LIMITACIONES DE LOS ENSAYOS CLÍNICOS Y DE LOS METAANÁLISIS

El metaanálisis es una herramienta ciertamente prometedora en la investigación y en la evaluación de las tecnologías médicas, pero de ninguna manera, hoy día,

es una herramienta infalible. El sesgo de publicación es un factor que hay que tener en cuenta, aunque se trate de un sesgo a veces difícil de medir.

Habitualmente el sesgo se produce por la no publicación de los estudios que dan resultados negativos y, por tanto, el efecto observado en el metaanálisis puede estar sobrevalorado. Cuando se trata de un tema de actualidad o de máximo interés, como puede ser el tratamiento hipolipemiente en prevención primaria y secundaria del infarto de miocardio, o cuando se trata de ensayos clínicos que incluyen un elevado número de pacientes, lo más probable es que se publiquen los ensayos clínicos sean o no los resultados positivos. Ahora bien, cuando se trata de temas de investigación con un interés secundario o que simplemente «no están de moda» o ensayos de pequeño tamaño, los posibles metaanálisis son más susceptibles al sesgo de publicación por no incluir los ensayos que daban resultados negativos. De hecho existe evidencia de que los ensayos clínicos de pequeño tamaño están más afectados por el sesgo de publicación que los ensayos clínicos de gran tamaño⁴. Y, además, dado que se trata de una muestra pequeña, la magnitud del efecto observado ha de ser grande para que sea estadísticamente significativo, lo que induce todavía más a sobrevalorar el efecto. Aunque inicialmente se pensó que las revistas médicas podían ser parcialmente responsables del sesgo de publicación, ya que tenderían a publicar sólo estudios con resultados significativos, actualmente se sabe que son los propios autores o la combinación de ambos, al considerar que si no hay resultados significativos los estudios «no son interesantes». En este sentido, el esfuerzo que está haciendo la Colaboración Cochrane⁵ con la actualización de bases de datos de ensayos clínicos y algunas revistas biomédicas como el *British Medical Journal*⁶ invitando a los investigadores a enviar todos aquellos ensayos clínicos no publicados puede contribuir a minimizar el sesgo de publicación.

Otro aspecto que se ha comentado respecto a los ensayos clínicos y a los metaanálisis es el fenómeno del «deseo de complacer» («willingness to please»)⁷, fenómeno complejo y que se puede producir desde las mejores y más honestas intenciones. Un ejemplo de esta complacencia se intuyó en un metaanálisis que agrupaba ensayos clínicos en los que se comparaba la utilización de heparina de bajo peso molecular con la heparina clásica⁸. En este metaanálisis, en los ensayos con peor puntuación cualitativa se observaron menos trombosis y menos hemorragias y en los ensayos de mejor calidad y publicados en revistas de más prestigio se detectaron menos trombosis pero más hemorragias con la heparina de bajo peso molecular. El resultado combinado, después del metaanálisis, fue que globalmente el nuevo tratamiento era más eficaz y producía menos efectos secundarios que la heparina clásica. Después de la publicación de este metaanálisis se publicó un megaensayo⁹, en el que 3.809 pacientes sometidos a

cirugía abdominal fueron aleatorizados a un tipo de heparina u otra, no encontrando diferencias en el riesgo de trombosis, pero sí una reducción, aunque no significativa, en el riesgo de hemorragias mayores con la heparina de bajo peso molecular (riesgo relativo [RR] 0,77; intervalo de confianza [IC] del 95%, 0,56-1,04). Este hecho pone en evidencia las discrepancias entre los metaanálisis y los megaensayos que ya se comentarán posteriormente, pero también pone en evidencia que todos los ensayos clínicos que se habían hecho hasta el metaanálisis (incluido éste) ofrecieron unos resultados que «complacían» la nueva intervención.

Este fenómeno también se ha observado en los ensayos clínicos realizados con homeopatía e incluso en un metaanálisis de los mismos, en el que los investigadores demostraron cierta eficacia de la homeopatía para algunas enfermedades. Se ha mencionado¹⁰ que un ensayo clínico con «diluciones infinitas» como es la homeopatía, comparado con una solución placebo, es como comparar dos placebos y, por tanto, los resultados positivos de la homeopatía sólo pueden ser consecuencia del azar. Se sospecha que estos resultados positivos a favor de las diluciones infinitas de la homeopatía podrían ser explicados en parte por el deseo de complacer la intervención. Se trataría, pues, de un fenómeno que actúa por unos mecanismos difíciles de explicar sobre todo delante de un ensayo clínico con una aleatorización y enmascaramiento correctos. Otras veces se pueden detectar defectos en la asignación aleatoria a un grupo u otro pudiendo influir en los resultados. Schulz et al¹¹ demostraron empíricamente que los ensayos en los que la secuencia de asignación se había ocultado de una forma inadecuada producían estimaciones más elevadas en los efectos de tratamiento en comparación con los ensayos en los que los autores describían una ocultación adecuada de la asignación. Otra limitación de muchos metaanálisis y ensayos clínicos es cuando no se trata de ensayos aleatorizados con control placebo, sino que se trata de ensayos aleatorizados con un grupo control que representa la práctica habitual. En estos casos, es difícil discernir cuál es el efecto real del fármaco a ensayar. Por ejemplo, en un metaanálisis sobre el efecto de la amiodarona en reducir la mortalidad en pacientes con infarto de miocardio e insuficiencia cardíaca¹², se agruparon los datos de 8 ensayos clínicos en los que comparaban la amiodarona con placebo o la amiodarona con la atención habitual que recibían los pacientes. Los resultados observados fueron que la amiodarona producía una modesta aunque significativa reducción absoluta de la mortalidad total (1,4%) y de las arritmias o muerte súbita combinadas (1,7%). En un análisis de subgrupos realizado según si el control era placebo o si el control era atención habitual observaron una reducción de la mortalidad total significativamente mayor en el segundo grupo, lo cual es difícil de explicar. Los betabloqueantes y los inhibidores de la en-

zima conversiva de la angiotensina (IECA) son fármacos frecuentemente utilizados en estos pacientes, lo que no está claro es si estos fármacos fueron utilizados de una manera consistente por estos pacientes, y si había algún tipo de interacción con la amiodarona.

Para mejorar éstas y otras limitaciones de los ensayos clínicos y los metaanálisis se han publicado recientemente¹³ unas Normas de Calidad Consolidadas para la publicación de ensayos clínicos (CONSORT) que contienen 21 puntos que corresponden principalmente a los apartados de métodos, resultados y discusión de un ensayo clínico e identifican elementos clave de información necesarios para evaluar la validez interna y externa de la publicación, y un diagrama de flujo que aporta información acerca del proceso seguido por los pacientes a lo largo de un ensayo clínico con un diseño de dos grupos paralelos que es, tal vez, el tipo de ensayo que se presenta con más frecuencia; también se ha propuesto recientemente una guía para el diseño, realización e interpretación de los metaanálisis¹⁴ que finalmente deberían cumplir unos estándares mínimos, al igual que ocurriría con un único ensayo clínico bien diseñado. En la **tabla 1** se exponen las características más importantes de esta propuesta. Concretamente, destaca la necesidad de desarrollar un protocolo prospectivo, definiendo las hipótesis primarias y secundarias que se pretenden contrastar, la utilización de ensayos clínicos aleatorizados adecuados, de realizar un análisis por intención de tratar (incluyendo en el análisis todos los pacientes manteniéndolos en los grupos a los que habían sido asignados aleatoriamente, independientemente de cómo acabaron en el ensayo) y basar las conclusiones en la globalidad de la información, incluyendo las variables de resultados más relevantes, todos los pacientes aleatorizados en los ensayos y de todos los ensayos clínicos publicados o no publicados. Se da mucha importancia al carácter prospectivo en la realización del metaanálisis, es decir, a la elaboración de un protocolo previo, la formulación de hipótesis primarias o secundarias, la definición de los criterios de inclusión y exclusión, la elección del tratamiento específico de interés y de las variables de resultados principales y secundarias, la utilización de datos individuales originales, definiendo los datos que se desean extraer y el tipo de análisis que se piensa realizar. No es frecuente encontrar en la bibliografía metaanálisis que cumplan con estas características, ya que el esfuerzo a realizar es mucho mayor. Ejemplos de metaanálisis con diseño prospectivo serían el Fibrinolytic Therapy Trialists' Collaboration¹⁵, un metaanálisis de ensayos con hipolipemiantes¹⁶ o un metaanálisis todavía en fase del protocolo, que revisa los grandes ensayos clínicos en curso en el tratamiento de la hipertensión¹⁷. También es importante saber si un metaanálisis es concluyente o bien genera nuevas hipótesis. En el caso de los metaanálisis de los hipolipemiantes¹⁶ se observó que en algunos de los ensayos se

TABLA 1
Guía para el diseño, realización
e interpretación del metaanálisis

Criterios cualitativos

1. Desarrollar un protocolo prospectivo definiendo las hipótesis primarias y secundarias que se quieren contrastar
 - Formular una pregunta que sea médicamente relevante y biológicamente plausible
 - Definir la población a estudiar y los subgrupos relevantes
 - Definir los tratamientos a incluir
 - Definir las variables de resultados de interés
 - Identificar los ensayos prospectivamente (idealmente a través de registros prospectivos)
 - Invitar a los investigadores de los ensayos clínicos a colaborar, obteniendo si es posible los datos de los pacientes
2. Utilizar sólo ensayos clínicos apropiados
 - Revisar la integridad de la aleatorización
3. Utilizar un análisis de intención de tratar y basar las conclusiones en la totalidad de la información, incluyendo
 - Todos los resultados relevantes registrados en el ensayo
 - Todos los pacientes aleatorizados
 - Todos los ensayos (búsqueda exhaustiva de la bibliografía incluyendo información de ensayos no publicados)
4. Revisar aspectos clave de la calidad y la integridad de los datos

Criterios cuantitativos

- Calcular el tamaño de muestra óptimo, basado en suposiciones realistas
- Utilizar márgenes de monitorización para sopesar la nueva evidencia haciendo actualizaciones periódicas
- Utilizar métodos estadísticos válidos para poder combinar los datos

Interpretación

- ¿Es el metaanálisis prospectivo o retrospectivo?
- ¿Es el metaanálisis concluyente o genera hipótesis?
- Valorar la potencia (error beta) del metaanálisis de la misma manera que se hace en un ensayo clínico
- Considerar la posibilidad de un error alfa (observar diferencias cuando en realidad no existen), particularmente cuando existen pocos eventos
- Considerar la posibilidad del sesgo de publicación

Adaptada de Pogue J y Yusuf S¹⁴.

producía una reducción del riesgo de accidente cerebrovascular. A raíz de este metaanálisis se generó una hipótesis de trabajo, que se contrastó en posteriores metaanálisis, siendo éstos ya concluyentes, confirmando que los hipolipemiantes, además de reducir el riesgo de enfermedad coronaria también reducen el riesgo de accidente cerebrovascular^{18,19}.

En general se aconseja²⁰ que el lector crítico de metaanálisis sea conservador y concluya que el verdadero efecto de la intervención es más pequeño que el estimador observado porque no se puede garantizar la no existencia de sesgos, aunque se trate de un metaanálisis realizado con la mayor pulcritud.

APLICACIÓN DE LOS RESULTADOS DEL METAANÁLISIS AL PACIENTE INDIVIDUAL

La aplicación de los resultados de los metaanálisis a un paciente individual requiere la aceptación de que la población incluida en los ensayos clínicos es un grupo apropiado de referencia para el paciente en cuestión. Dado que un único resultado global del estimador del efecto puede agrupar poblaciones muy diversas, una alternativa es buscar aquellos subgrupos con características muy parecidas a las de nuestro paciente. Sin embargo, el emparejamiento de las características de nuestro paciente con las de algún subgrupo nos puede llevar a interpretaciones erróneas. Por ejemplo, imaginemos a un médico que trabaja en Alemania y quiere saber hasta qué punto los resultados del metaanálisis sobre betabloqueantes e infarto de miocardio –con una clara reducción de la mortalidad– puede aplicarlos a los pacientes de su consulta. El médico revisa los ensayos clínicos incluidos en el metaanálisis y observa que el único estudio en el que se incluyeron un número importante de pacientes alemanes era el European Infarction Study²¹. Precisamente este estudio fue de los pocos que demostró que no existía ningún beneficio, e incluso se observó un aumento de la mortalidad (no estadísticamente significativa) en los pacientes tratados con betabloqueantes. ¿Cómo deberían tratarse entonces los pacientes en Alemania? Seguramente, es altamente improbable que los pacientes alemanes que han padecido un infarto no se beneficiaran del tratamiento betabloqueante al igual que ocurría en pacientes de otros países europeos. En este estudio en concreto se utilizó el oxprenolol, fármaco con actividad simpaticomimética intrínseca, aceptándose que este grupo de fármacos no ofrece un beneficio claro. Se ha afirmado que la aplicación de un estimador global de efecto del metaanálisis se basa en la premisa de que las inconsistencias observadas entre los ensayos en los subgrupos se deben al azar, e implicaría desviar los estimadores puntuales de estos subgrupos al gran valor medio. El grado de ajuste de un valor individual hacia ese valor medio dependerá de su precisión y, por tanto, del tamaño de la submuestra. El European Infarction Study representaba una mínima contribución dentro de la totalidad de la evidencia y, por tanto, el estimador observado necesitaría ajustarse considerablemente en la dirección del estimador global.

Un factor que, sin duda, contribuye a aplicar inadecuadamente los resultados de los análisis de subgrupos es la afición que existe por parte de los autores de un artículo de metaanálisis en concentrarse a veces, incluso más, en los análisis de subgrupos que no en los resultados globales. Un ejemplo de esto se observó en un estudio²² en el que se reanalizaban los datos del Beta-Blocker Heart Attack Trial²³, donde globalmente los pacientes que habían padecido un infarto y estaban en

tratamiento con propranolol tenían una reducción de la mortalidad de un 27%. En este estudio se observó que los resultados de un tercio de los centros (11 de 31) eran discordantes con los de la globalidad, observándose incluso un aumento del riesgo de mortalidad del 33% en esos centros (no estadísticamente significativo). Se llegó a sugerir²² incluso que se cambiaran las recomendaciones del tratamiento betabloqueante según las características de los pacientes de los centros en los que se observó un beneficio (pacientes con historia de infarto de miocardio, con insuficiencia cardíaca y de raza no blanca). Estos resultados, sin embargo, son muy probablemente consecuencia del azar y cualquier conclusión puede ser espúrea, pudiéndose producir resultados contradictorios que lo único que aportan es confusión en la decisión terapéutica²⁴. Podría parecer que un efecto demostrado para el conjunto de la población de estudio con un tamaño de muestra amplio se podría analizar fácilmente en diferentes subgrupos de pacientes (según la edad, sexo, antecedentes, enfermedades concomitantes, etc.) y acercarnos así a la predicción del efecto en el paciente individual. Esta observación aparentemente obvia es una falacia y el análisis de subgrupos es a menudo problemático²⁵⁻²⁷. Un análisis retrospectivo de subgrupos definido después del estudio no es equivalente a lo que hubiera sido la aleatorización independiente de los pacientes de cada subgrupo. Por tanto, el análisis de subgrupos hecho *post hoc*, cuando ya se dispone de la estimación del efecto sobre la población global, y más cuando se empieza a dividir y subdividir la población para hacer comparaciones múltiples en función de cómo salieron los resultados, no tiene prácticamente validez. Esta manera de actuar, que no es excepcional, se ha denominado «tortura de los datos» («Cuando se torturan los datos, éstos dicen aquello que al final se quería que dijeran»).

Para que el análisis de subgrupos sea fiable se exigen una serie de condiciones, resumiéndose las más importante en la **tabla 2**. Cuando el análisis de subgrupos no cumpla alguna de ellas, hay que poner en duda la validez de sus resultados. Posiblemente, la condición más importante es la definición de los subgrupos en la fase de diseño del estudio, cosa que puede permitir una previsión adecuada de los requisitos necesarios para la validez, por ejemplo, cuál tiene que ser el tamaño adecuado de la muestra de cada subgrupo.

También se ha afirmado que para realizar una estimación del efecto en determinados subgrupos puede ser más orientativo un juicio crítico adecuado sobre el resultado global del estudio, que incluya una evaluación de su metodología y de la plausibilidad biológica del efecto en los subgrupos correspondientes, que un análisis formal de subgrupos, en especial si es retrospectivo.

Estas consideraciones no representan que el análisis retrospectivo de subgrupos sea totalmente carente de valor. Su principal utilidad es, como ya se ha comentado anteriormente, sugerir o generar hipótesis, que des-

TABLA 2
Criterios de validez del análisis de subgrupos

Cuando la hipótesis específica se había planteado ya en el diseño del estudio
Cuando la magnitud del efecto observado es grande (RRR superior a un 25%)
Cuando el efecto del tratamiento es altamente significativo (cuanto más bajo es el valor de la p más creíble será la diferencia)
Cuando los resultados son consistentes con los de otros estudios
Cuando exista una evidencia indirecta que vaya a favor de los resultados encontrados (plausibilidad biológica)

RRR: reducción del riesgo relativo.

pues deberían ser demostradas en estudios específicos. Un ejemplo reciente de ensayo clínico, en el que realizan análisis de subgrupos de una manera correcta, es el primer gran ensayo clínico que evaluaba el efecto de la digital en el tratamiento en pacientes con insuficiencia cardíaca y ritmo sinusal²⁸. Los resultados de este ensayo no demostraron un efecto beneficioso sobre la mortalidad total (RR 0,99; IC del 95%, 0,91-1,07) y una pequeña reducción cuando se combinaban las variables de resultado de hospitalización o muerte por empeoramiento de la insuficiencia cardíaca (RR 0,75; IC del 95%, 0,69-0,82). Estos resultados no son igualmente generalizables a toda la población del estudio ya que se observó un mayor beneficio —tanto en términos absolutos como relativos— respecto a la última variable de resultado combinada en los pacientes de riesgo alto, con fracciones de eyección por debajo de 0,25 o cardiomegalia y clase funcional de la New York Heart Association (NYHA) entre III y IV. En este estudio se habían definido previamente en el diseño del estudio los subgrupos que posteriormente se analizaron, por lo que es de suponer que se formularon hipótesis específicas para cada análisis de subgrupos. Todas estas consideraciones sobre el análisis de subgrupos se aplican no sólo a los ensayos clínicos y metaanálisis sino también a los megaensayos, que se comentarán posteriormente.

En la toma de decisiones en el paciente individual, los estudios observacionales de efectividad pueden aportar información complementaria útil, sobre todo porque se trata de estudios en los que se incluyen pacientes no seleccionados con enfermedades concomitantes o con ciertas características que hacen que sistemáticamente sean excluidos de los ensayos clínicos. Recientemente se ha publicado un estudio realizado a partir de un registro de los pacientes de Medicare²⁹, en el que recopiló información clínica de 201.752 pacientes diagnosticados de infarto de miocardio entre febrero de 1994 y julio de 1995. Utilizando un análisis de regresión para controlar todas las variables que pudieran influir en el pronóstico, se comparó la mortalidad a los 2 años del infarto entre los pacientes que estaban en tratamiento con betabloqueantes y los que no esta-

ban con este tratamiento. Para aquellos pacientes que no habían padecido complicaciones se observó una reducción de la mortalidad en un 40%. También se observó beneficio para aquellos grupos de pacientes en los que habitualmente los betabloqueantes se consideran contraindicados, como son los pacientes con insuficiencia cardíaca y pacientes ancianos, pacientes con infarto sin onda Q y aquellos con enfermedad pulmonar obstructiva crónica. En el «Editorial» que acompaña al artículo³⁰ se comenta, sin embargo, que al analizar aquellas posibles contraindicaciones sin tener en cuenta la severidad de la enfermedad, como la enfermedad pulmonar obstructiva crónica o la insuficiencia cardíaca, los resultados pueden estar sesgados hacia un resultado beneficioso para estos grupos. Aunque también se comenta que la gran virtud de estos estudios observacionales es que permiten evaluar la efectividad terapéutica, es decir, los resultados observados en un mundo más real que el de los ensayos clínicos.

GENERALIZACIÓN DE LOS RESULTADOS

La generalización de los resultados, también conocida como la validez externa, incumbe tanto a ensayos clínicos y metaanálisis como a megaensayos. El problema de generalización de los resultados tradicionalmente se ha relacionado con el hecho de extrapolar los mismos a grupos no incluidos en el ensayo clínico como pudieran ser las mujeres en el caso de haberse realizado en varones, o extrapolarse a personas mayores de 65 años en el caso de haberse centrado en adultos jóvenes. En estos casos ha habido y habrá controversia entre diferentes grupos de investigadores, y extrapolarlo más o menos a otras poblaciones dependerá de los resultados de otros estudios epidemiológicos no experimentales, la consistencia entre los mismos y su validez interna. Un ejemplo de discrepancia abierta en la bibliografía médica ocurre en la prevención primaria de la enfermedad coronaria, concretamente en el tema del cribado del colesterol sérico en pacientes asintomáticos y en el tratamiento hipolipemiente. Así el American College of Physicians³¹ recomienda el cribado sólo en varones entre 35 y 65 años y en mujeres entre 45 y 65 años, utilizando el colesterol total sin considerar el colesterol-HDL, el National Cholesterol Educational Program/AHA³² recomienda incluir tanto a varones como a mujeres desde los 35 años, sin límite de edad superior, y además del colesterol total incluyen el colesterol-HDL. Respecto al tratamiento hipolipemiente, existen los defensores de los valores de colesterol y fracciones lipídicas exclusivamente para empezar el tratamiento farmacológico, y los que abogan por un abordaje más integral utilizando el riesgo coronario de padecer un evento coronario a los 10 años como elemento decisorio³³. Al trasladar la evidencia a la práctica, además de las consideraciones comentadas previamente, también hay que tener en cuenta la magnitud de

la enfermedad en el entorno concreto en el que se va aplicar esa evidencia. Por ejemplo, en un intento de simulación de la aplicación del estudio WOSCOPS (estudio de prevención primaria de infarto de miocardio realizado en Escocia) en nuestro medio³⁴, para una misma reducción del riesgo relativo del 31% la reducción del riesgo absoluto fue del 0,62%, mientras que en Escocia fue del 2,4%. Esto es así porque la incidencia de la enfermedad coronaria estimada en nuestro medio es del 0,4%, mientras que la incidencia observada en el grupo placebo del estudio WOSCOPS³⁵ era del 1,6%.

DEL METAANÁLISIS AL MEGAENSAYO

¿Es la muestra de un metaanálisis lo suficientemente amplia como para obtener conclusiones válidas? Por ejemplo, en los ensayos clínicos con trombolíticos, el beneficio del tratamiento raramente supera una reducción relativa del riesgo del 25%. Para poder detectar estas diferencias se necesitan observar al menos 800 eventos. Si el número de eventos observados en un metaanálisis es alrededor de esa magnitud y la significación estadística es muy alta ($p < 0,0001$), entonces existen pocas dudas sobre la eficacia del tratamiento. Por contra, si el número de eventos es moderado (menos de 200) o la significación estadística es modesta ($p < 0,05$), el resultado es más cuestionable, y se necesita realizar un ensayo clínico de gran tamaño.

¿QUÉ ES UN MEGAENSAYO?

Los megaensayos son ensayos clínicos aleatorizados con un tamaño de muestra muy grande, necesaria para demostrar un efecto de magnitud relativamente pequeño. En cardiología, su principal objetivo ha sido la evaluación de fármacos en el infarto agudo de miocardio y, en este sentido, el estudio GISSI-1, que evaluó la eficacia de la estreptocinasa en el infarto agudo de miocardio con una muestra de 14.000 enfermos, marcó un hito en la historia de la terapéutica³⁶. No existe una definición precisa de cuál es el número de pacientes que definen un megaensayo, pero se consideran así los ensayos clínicos con más de 10.000 enfermos; se han llegado a incluir hasta casi 60.000 en el estudio ISIS-4³⁷. Quizá parece excesivamente entusiasta la opinión de Topol y Califf, impulsores de los megaensayos GUSTO, que definen como «miniensayos» aquellos ensayos aleatorizados que incluyen hasta 3.000 pacientes³⁸. Las características fundamentales que diferencian los ensayos clínicos convencionales de los megaensayos son las siguientes:

1. Los ensayos clínicos pretenden estudiar a una población más bien definida, caracterizada por unos criterios de inclusión y exclusión muy estrictos y un riguroso control de los tratamientos concomitantes. El protocolo puede ser complejo, y se evalúan diferentes

variables de resultado, incluida a veces la mortalidad: variables clínicas, de estudios funcionales o de laboratorio, o de calidad de vida.

2. Los megaensayos seleccionan a una población muy amplia y heterogénea, con unos criterios de inclusión y exclusión poco estrictos (p. ej., infarto de miocardio o sospecha de infarto de menos de 24 h de evolución, independientemente de su localización o complicaciones, de la edad y el sexo, enfermedades concomitantes u otros tratamientos). Esto, sin duda, permite incluir a un gran número de enfermos con una gran variedad de características, y procedentes de muchos centros de diferentes países. Esta característica tiene tres consecuencias principales: en primer lugar, la única variable de resultado que se puede determinar mejor en estas condiciones tan heterogéneas, virtualmente la única, es la mortalidad. Otras variables, todas ellas menos inequívocas que la mortalidad, serían de medición difícil o cuestionable. La segunda consecuencia es, quizá, la característica más importante de los megaensayos: debido al tamaño tan amplio de la muestra, sus resultados consiguen un alto grado de precisión (intervalos de confianza muy estrechos). Son estudios con una potencia estadística suficiente para detectar diferencias estadísticamente significativas, aunque la magnitud del efecto observado sea pequeña. En tercer lugar, así como en los estudios más pequeños sólo podemos afirmar que el efecto observado se produce en una población de características limitadas, en un megaensayo sabemos que el efecto tiene lugar en una población amplia y heterogénea. Esto hace que las características de la población de los megaensayos se aproximen a las de los estudios de efectividad.

APLICACIÓN DE LOS RESULTADOS DE LOS MEGAENSAYOS AL PACIENTE INDIVIDUAL

Al igual que ocurre con los ensayos clínicos o los metaanálisis, aquí también nos hemos de preguntar si el efecto demostrado en la población de estudio puede ser aplicado a un paciente concreto, por ejemplo, si la reducción de mortalidad observada con un fármaco trombolítico puede ser aplicada al paciente que estamos tratando. La respuesta a esta pregunta es, hasta cierto punto, paradójica, ya que resulta que se puede demostrar que un fármaco trombolítico es eficaz para reducir la mortalidad en general pero no nos dice si será útil para el paciente concreto que estamos tratando.

En un trabajo reciente se ha llegado a hacer una crítica demoledora del concepto sobre el que se basa el megaensayo, y afirma que la heterogeneidad de su población hace que ésta sea una población abstracta, que no tenga una existencia real y no se pueda, por tanto, inferir ningún efecto sobre el paciente individual. El autor afirma³⁹: «La estimación proveniente de un estudio que incluya individuos heterogéneos (p. ej., un megaensayo) no nos dice *nada* de la experiencia de cada

uno de ellos, ni en lo que se refiere a los incluidos en el estudio (*validez interna*) ni a los observados en otras circunstancias (*validez externa*)... En los megaensayos se ha conseguido un aumento de la precisión al precio de una reducción de la validez: intervalos de confianza estrechos en torno a unas estimaciones inválidas.» Esta crítica, aunque no le falte fundamento, probablemente es excesiva y no anula el valor del megaensayo, ya que su finalidad no es pronosticar la respuesta individual sino definir un efecto en una población.

Ésta es una característica de los ensayos clínicos en general, y ya fue destacada hace muchos años por Sir Austin Bradford Hill, uno de los introductores y gran defensor del ensayo clínico, quien afirmó: «El ensayo clínico puede demostrar, sin ninguna duda, que el tratamiento A es, como media, superior al tratamiento B... No contesta la pregunta del médico: ¿cuál será el resultado si este fármaco se administra a este paciente en concreto?»⁴⁰. Este dilema alcanza su máxima expresión con el megaensayo.

Las mismas limitaciones que se han comentado para el análisis de subgrupos del metaanálisis pueden aplicarse a los megaensayos.

Un ejemplo pintoresco para ilustrar lo que puede pasar cuando se analizan diferentes subgrupos sin las condiciones anteriormente mencionadas en la **tabla 2**, se produjo intencionadamente en el análisis del megaensayo ISIS-2, en el que se compararon la aspirina, la estreptocinasa o ambas con placebo en el tratamiento del infarto agudo de miocardio. En un análisis retrospectivo de subgrupos se observó una ausencia de efecto de la aspirina comparada con el placebo en los pacientes de los signos del Zodíaco Libra y Géminis, y una diferencia altamente significativa para el resto de signos⁴¹. Este resultado ilustra la falacia de la técnica de análisis, a pesar del tamaño de la muestra del megaensayo (más de 17.000 enfermos) que aparentemente permitiría este tipo de análisis.

La magnitud del efecto observado en el megaensayo es importante para poder evaluar su aplicabilidad. Aunque en teoría no tiene por qué haber una relación entre la validez externa y la magnitud (un efecto real puede ser grande o pequeño), lo cierto es que en la práctica los efectos de magnitud grande son más fácilmente interpretables, tanto en lo que se refiere a la validez externa, como en el análisis de subgrupos. Si no existe la sospecha de una apreciable ausencia de efecto en subgrupos específicos de suficiente tamaño y es plausible biológicamente el efecto observado en los subgrupos específicos, en la práctica se asume que la intervención evaluada también es efectiva en los mismos. Éste es el tipo de razonamiento que se hace habitualmente en la aplicación de los megaensayos en la realidad cotidiana, más que hacer un análisis de subgrupos formal. Por ejemplo, el beneficio de la trombolisis, calculado a partir de los megaensayos, representa un efecto no pequeño⁴²: una reducción del riesgo de

mortalidad entre el 18 y el 25%, siendo el número necesario de pacientes a tratar para evitar un caso de muerte (NNT) entre 35 y 45, aproximadamente. Esto representa un efecto terapéutico de suficiente magnitud como para suponer que se ejerce sobre el conjunto de la población con infarto de miocardio y sus subgrupos, a pesar de que el riesgo individual sea poco previsible. Como se sabe, las excepciones son los pacientes con depresión del segmento ST y los pacientes con infarto de más de 12 h de duración, que no parece, examinando los resultados globales de los ensayos, que se beneficien del tratamiento (no obstante, precisamente a causa de las incertidumbres del análisis retrospectivo de subgrupos, no existe una evidencia totalmente concluyente sobre esta falta de efecto). Pero si el efecto es pequeño las cosas se complican, ya que entonces se suscitan muchas más dudas sobre la aplicabilidad de los resultados al paciente individual. Un ejemplo de este hecho es el efecto de los inhibidores de la enzima conversiva de la angiotensina (IECA), administrada durante las primeras 24 h después de un infarto agudo de miocardio. En el estudio CONSENSUS-2, que incluyó a 6.000 pacientes⁴³ no se demostró ningún efecto del enalapril en esta situación. En cambio, en dos megaensayos (GISSI-3 e ISIS-4)^{37,44} se comprobó un efecto de reducción de la mortalidad con lisinopril y captopril, respectivamente, ejemplo que ilustra la capacidad de los megaensayos para demostrar efectos de magnitud de significación estadística que no son evidenciables en otros estudios. Este efecto resultó claramente significativo, pero pequeño: correspondería tratar a unos 125-200 enfermos para prevenir una sola muerte. Con un efecto de magnitud tan pequeño es comprensible que se planteen dudas sobre la aplicabilidad a todos los pacientes. Se ha puesto en duda, por ejemplo, que esté justificado tratar a todos los pacientes, y que el mayor efecto pudiera tener lugar sólo en un subgrupo de pacientes identificable por otros medios (dilatación o disfunción ventricular izquierda)⁴⁵. Se ha afirmado también que, en la práctica asistencial cotidiana, fuera del contexto de un ensayo clínico (en el que la atención y el control de la medicación administrada pueden ser diferentes), el beneficio neto de administrar el fármaco a todos los pacientes puede verse sobrepasado por sus efectos indeseables⁴⁶.

Otro ejemplo en el que se combinan la noción de magnitud del efecto y el juicio individual en la evaluación de subgrupos es la consideración de los enfermos con riesgo bajo espontáneo (p. ej., un primer infarto inferior poco extenso y sin complicaciones en un paciente joven). Aunque la reducción relativa del riesgo se mantenga idéntica a la observada en la población global del estudio, su traducción en reducción absoluta puede ser lo suficientemente pequeña como para que algunos médicos crean inaceptable el tratamiento trombolítico delante del riesgo, bajo pero real, de hemorragia cerebral⁴⁷.

Estos ejemplos comentados ilustran cómo un efecto de escasa magnitud observado en la población heterogénea de un megaensayo puede plantear dudas al médico clínico, delante de la incertidumbre de cuáles son realmente los enfermos que se benefician del tratamiento. Ésta es la pregunta que el megaensayo precisamente no contesta.

DISCREPANCIAS ENTRE METAANÁLISIS Y MEGAENSAYOS

No es un hecho infrecuente observar resultados contradictorios entre los metaanálisis y los megaensayos, y antes ya se ha comentado el ejemplo del tratamiento anticoagulante. En un artículo reciente de revisión⁴⁷ en el que se comparaban los metaanálisis y los ensayos clínicos amplios (considerados éstos como los que incluían a más de 1.000 pacientes), se observó que ambos métodos tienden a estar en desacuerdo en un 10-23% de las veces. Quizá la discrepancia más sonora fue la del metaanálisis que refirió que administrar magnesio intravenoso a los pacientes que habían padecido un infarto de miocardio era beneficioso⁴⁸, cuando un megaensayo posterior en donde participaron más de 58.000 pacientes³⁷ no observó ningún beneficio. Las deficiencias de este metaanálisis radicaban en un sesgo de publicación (se incluyeron fundamentalmente los ensayos clínicos que daban resultados positivos), en la insuficiencia metodológica de los ensayos clínicos pequeños y en la heterogeneidad clínica.

Otras veces se ha observado, mediante un metaanálisis acumulado, que la realización de más ensayos clínicos o megaensayos era innecesaria. Por ejemplo, si se hubiera combinado la evidencia existente en 1973 sobre el tratamiento de estreptocinasa intravenosa del infarto de miocardio, ya se hubiera observado una reducción estadísticamente significativa ($p = 0,01$) de la mortalidad (en ese momento se habían aleatorizado 2.432 pacientes de 8 ensayos clínicos). En el año 1986, justo antes de la publicación del primer megaensayo GISSI-I³⁶, la acumulación de la evidencia aportaba una reducción todavía más significativa ($p = 0,001$), y después de la publicación del estudio ISIS-2⁴¹ la p era de $< 10^{-15}$. Se ha afirmado que realizar más estudios cuando la evidencia es contundente es un ejercicio superfluo, costoso y no ético⁴⁰; aunque pudiera suceder, como hemos visto con el ejemplo de los anticoagulantes o el del tratamiento con magnesio, que los megaensayos contradigan la evidencia previa de los metaanálisis. Lo más razonable, antes de realizar el esfuerzo enorme que representa un ensayo clínico y ya no digamos un megaensayo, sería revisar concienzudamente la evidencia existente hasta la fecha, explorando todas las limitaciones de los ensayos clínicos y los megaensayos por un comité de expertos independiente, que fuera el que decidiera sobre la necesidad de realizar más estudios⁴⁹.

LÍNEAS DE INVESTIGACIÓN FUTURAS Y CONCLUSIONES

Parece evidente que los megaensayos han permitido la introducción de medidas efectivas en el tratamiento del infarto de miocardio. De hecho, en pocos años se ha conseguido que el mundo cardiológico viva pendiente de los megaensayos que se van completando antes de introducir cambios, ni que sean mínimos, en la práctica asistencial. Además, los megaensayos pueden haber representado una cierta revolución conceptual, al ser una de las maneras de obtener evidencia aleatorizada a gran escala, justificada para demostrar efectos de magnitud pequeña o mediana (la otra de las maneras es el metaanálisis), dejando de lado la exigencia clásica de los ensayos clínicos de demostrar el efecto en poblaciones bien definidas y con un control estricto de las terapéuticas concomitantes. Pero los megaensayos han suscitado críticas, y ya se han comentado algunas como la heterogeneidad de la población y su dificultad en aplicar los resultados individualmente. Esta crítica es más una señal de advertencia que una insuficiencia que anule su valor. Existen otras posibles críticas. Esta arma tan poderosa se ve limitada por su propia magnitud cuando se piensa en su aplicación futura. De hecho, la realización de un megaensayo comporta un esfuerzo inmenso de coordinación, logístico y económico para conocer prácticamente el efecto de un tratamiento sobre una única variable de resultado (la mortalidad), sin obtener información sobre otras variables de indudable valor clínico. Cualquier evaluación del efecto de otro fármaco (ni que sea emparentado con los ya ensayados) o del mismo fármaco sobre otra variable de resultado, puede requerir otro megaensayo con todo lo que representa. Esta situación puede expandirse *ad infinitum*. Por tanto, en el futuro se deberían definir muy bien las situaciones en las que el crecimiento y multiplicación de los fármacos de cada una de las familias donde se prevé un posible adelanto en la utilidad terapéutica (fibrinolíticos, nuevos antitrombóticos, IECA, hipolipemiantes, etc.) justifique el monumental uso de recursos humanos y materiales necesarios para obtener la evidencia que hemos comentado. ¿Es realmente necesario para el médico clínico y el gestor sanitario saber que una innovación terapéutica permita, pongamos por caso, salvar dos vidas más por mil enfermos que la estrategia ya existente, y saberlo mediante la realización de un megaensayo? ¿Para cuántos nuevos fármacos se deberá hacer? ¿Existen otras alternativas? Es posible que se necesite plantear una estrategia de amplio espectro, idealmente en el seno de organizaciones o instituciones de renombre internacional, poderosas e independientes, sin ánimo de lucro, que permitan una ordenación eficiente de las actividades de investigación con relación a las hipótesis de interés terapéutico susceptibles de ser evaluadas en megaensayos, y que eviten la redundancia y la confusión.

Otra pregunta podría ser qué aportan los megaensayos que no puedan aportar los metaanálisis, o cuándo éstos podrán suplir a los megaensayos en el futuro. No podemos ofrecer aquí una respuesta válida, y más sabiendo que el metaanálisis tiene todavía limitaciones en sus capacidades como hemos visto anteriormente. No obstante, como también se ha comentado antes, el metaanálisis acumulado antes de la realización del primer megaensayo GISSI-I ya ofrecía un resultado muy parecido al que después se observaría con el megaensayo, tanto en lo que se refiere a la magnitud del efecto como en el intervalo de confianza. Es decir, que desde este punto de vista el estudio GISSI-I fue innecesario. El metaanálisis es, por tanto, una estrategia a mejorar, que no debemos olvidar en un futuro que de otra manera, podría ser caótico. Concretamente, las recomendaciones descritas en la **tabla 1** pueden conducir en el futuro a la realización de menos metaanálisis, pero los que se realicen tendrán más garantías de calidad. Un papel fundamental del metaanálisis en el futuro será el de entender y predecir las discrepancias, y planificar nuevos ensayos clínicos teniendo *in mente* estos resultados. También sería deseable la realización de ensayos clínicos con la anticipación explícita de que en un futuro se incorporarán en un metaanálisis.

Tampoco debemos olvidar la información proveniente de los megaestudios de efectividad, como el comentado anteriormente²⁹. Es previsible y deseable que este tipo de estudios proliferen en un futuro no muy lejano, por los cada vez más asequibles sistemas de información clínica, y también, que se vayan sofisticando metodológicamente para mejorar su validez interna. No sólo nos ayudarán a conocer mejor lo que ocurre en la práctica asistencial, sino también pueden aportar una información muy valiosa para las decisiones clínicas individuales.

BIBLIOGRAFÍA

1. Guallar E, Damián J, Martín-Moreno JM. Metaanálisis y revisiones sistemáticas en cardiología. *Rev Esp Cardiol* 1997; 50: 345-354.
2. Baber NS, Evans DW, Howitt G. Multicentre post-infarction trial of propranolol in 49 hospitals in the United Kingdom, Italy and Yugoslavia. *Br Heart J* 1980; 44: 96-100.
3. Anti-Platelet Trialists' Collaboration. Secondary prevention of vascular disease by prolonged antiplatelet treatment. *Br Med J* 1988; 296: 320-331.
4. Begg CB, Berlin JA. Publication bias: a problem of interpreting medical data. *J R Statist Soc A* 1988; 151: 419-463.
5. Chalmers I. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann NY Acad Sci* 1993; 703: 156-163.
6. Smith R, Roberts I. An amnesty for unpublished trials. *Br Med J* 1997; 315: 622.
7. Vandembroucke JP. Observational research and evidence-based medicine: what should we teach young physicians? *J Clin Epidemiol* 1998; 51: 467-472.

8. Nurmohamed MT, Rosendaal FR, Büller HR, Dekker E, Hommes DW, Vanderbroucke JP et al. Low molecular weight heparin versus standard heparin in general and orthopedic surgery: a meta-analysis. *Lancet* 1992; 340: 152-156.
9. Kakkar VV, Cohen AT, Edmonson RA, Phillips MJ, Cooper DJ, Das SK et al. Low molecular weight versus standard heparin for prevention of venous thromboembolism after major abdominal surgery. The Thromboprophylaxis Collaborative Group. *Lancet* 1993; 341: 259-265.
10. Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV et al. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. *Lancet* 1997; 350: 834-843.
11. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273: 408-412.
12. Amiodarone Trials Meta-analysis Investigators. Effect of prophylactic amiodarone on mortality after acute myocardial infarction and in congestive heart failure: meta-analysis of individual data from 6500 patients in randomised trials. *Lancet* 1997; 350: 1.417-1.424.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. Special Communication: improving the quality of reporting of randomized controlled trials. The CONSORT Statement. *JAMA* 1996; 276: 637-639.
14. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomized controlled trials. *Lancet* 1998; 351: 47-52.
15. Fibrinolytic Therapy Trialists' Collaboration. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomized trials of more than 1,000 patients. *Lancet* 1994; 343: 311-322.
16. Byington RP, Jukema JW, Salonen JT, Pitt B, Brusckhe AV, Hoen H et al. Reduction in cardiovascular events of the pravastatin atherosclerosis intervention program. *Circulation* 1995; 92: 2.419-2.425.
17. World Health Organization-International Society of Hypertension Blood Pressure Lowering Treatment Trialists' Collaboration. Protocol for prospective collaborative overviews of major randomized trials of blood pressure lowering treatments. *J Hypertens* 1998; 16: 127-137.
18. Bucher HC, Griffith LE, Guyatt GH. Effect of HMGcoA reductase inhibitors on stroke. A meta-analysis of randomized, controlled trials. *Ann Intern Med* 1998; 128: 89-95.
19. Hebert PR, Gaziano JM, Chan KS, Hennekens CH. Cholesterol lowering with statin drugs, risk of stroke, and total mortality. An overview of randomized trials. *JAMA* 1997; 278: 313-321.
20. Yusuf S, Kitching AD. From journal to bedside: obtaining medically useful answers from a meta-analysis. *Evidence-Based Cardiovascular Medicine* 1997; 1: 85-86.
21. The European Infarction Study Group. European Infarction Study (EIS). A secondary prevention study with slow release oxprenolol after myocardial infarction: morbidity and mortality. *Eur Heart J* 1984; 5: 189-202.
22. Horwitz RI, Singer BH, Makuch RW, Viscoli CM. Can treatment that is helpful on average be harmful to some patients: a study of the conflicting information needs of clinical enquiry and drug regulation. *J Clin Epidemiol* 1996; 49: 395-400.
23. B-Blocker Heart Attack Trial Research Group. A randomized trial of propranolol in patients with acute myocardial infarction. I. Mortality results. *JAMA* 1982; 247: 1.707-1.714.
24. Smith GD, Egger M, Phillips AN. Beyond the grand mean. *Br Med J* 1997; 315: 1.610-1.614.
25. Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991; 266: 93-98.
26. Oxman AD, Guyatt GH. A consumer's guide to subgroup analysis. *Ann Intern Med* 1992; 116: 78-84.
27. Detsky AS, Naglie IG. Subgroup analysis: primary and secondary [editorial]. *ACP J Club* 1995; 122: A12-A14.
28. Digitalis Investigation Group. The effect of digoxin on mortality and morbidity in patients with heart failure. *N Engl J Med* 1997; 336: 525-533.
29. Gottlieb SS, McCarter RJ, Vogel RA. Effect of Beta-blockade on mortality among high risk and low risk patients after myocardial infarction. *N Engl J Med* 1998; 339: 489-497.
30. Radford M, Krumholz HM. Beta-blockers after myocardial infarction- for few patients, or many? *N Engl J Med* 1998; 339: 551-553.
31. American College of Physicians. Guidelines for Using Serum Cholesterol, High-Density Lipoprotein Cholesterol, and Triglyceride Levels as Screening Tests for Preventing Coronary Heart Disease in Adults. Part I. *Ann Intern Med* 1996; 124: 515-517.
32. National Cholesterol Program. Second report of the Expert Panel on detection, evaluation, and treatment of high cholesterol in adults (Adult Treatment Panel II). *Circulation* 1994; 89: 1.329-1.344.
33. Prevention of coronary heart disease in clinical practice. Recommendations of the Second Joint Task Force of European and Other Societies on Coronary Prevention. *Eur Heart J* 1998; 18: 1.434-1.503.
34. Brotons Cuixart C. Implicaciones en la práctica clínica de los megaestudios sobre prevención primaria con estatinas. *Rev Esp Cardiol* 1997; 50: 295-299.
35. Sheperd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, Macfarlane PW et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* 1995; 333: 1.301-1.307.
36. Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *Lancet* 1986; 1: 397-401.
37. ISIS-4 Collaborative Group. A randomized factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 1995; 345: 669-685.
38. Topol EJ, Califf RM. Answers to complex questions cannot be derived from «simple» trials. *Br Heart J* 1992; 68: 348-351.
39. Charlton BG. The scope and nature of epidemiology. *J Clin Epidemiol* 1996; 49: 623-626.
40. Feinstein AR. An additional basic science for clinical medicine: II. The limitations of randomized trials. *Ann Intern Med* 1983; 99: 544-550.
41. ISIS-2 Collaborative Group. Randomized trial of intravenous streptokinase, oral aspirin, both or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988; 2: 349-360.
42. Collins R, Peto R, Baigent C, Sleight P. Aspirin, heparin and fibrinolytic therapy in suspected acute myocardial infarction. *N Engl J Med* 1997; 336: 847-860.
43. Swedberg K, Held P, Kjekshus J, Rassmussen K, Rydén L, Wedel H, on behalf of the CONSENSUS II Study Group. Effects of the early administration of enalapril on mortality in patients with acute myocardial infarction. *N Engl J Med* 1992; 327: 678-684.
44. Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico. GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction. *Lancet* 1994; 343: 1.115-1.122.
45. Julian DG. Should ACE inhibitors be administered to all patients after acute myocardial infarction? A (cautious) negative response. *Eur Heart J* 1995; 16 (Supl E): 44-45.
46. Tobé TJM. Is thrombolytic therapy in acute inferior myocardial infarction really better than conventional treatment? *Br Heart J* 1995; 73: 108-109.
47. Ioannidis JPA, Cappelleri JC, Lau J. Issues in Comparison between meta-analysis and large trials. *JAMA* 1998; 279: 1.089-1.093.
48. Woods KL, Fletcher S, Roffe C, Haider Y. Intravenous magnesium sulphate in suspected acute myocardial infarction: results of the second Leicester Intravenous Magnesium Intervention Trial (LIMIT-2). *Lancet* 1992; 339: 1.553-1.558.
49. Murphy DJ, Povar GJ, Pawlson LG. Setting limits in clinical medicine. *Arch Intern Med* 1994; 154: 505-512.